# Alex Fields

## BAN 502

### Module 2 - Assignment 2

#### Task 2 - Q&A

It seems that “temp” and “atemp” are the highest correlated variables in relation to “Count” at (0.3).

#### Task 3 - Q&A

When looking at the box plots in figure 3, I noticed how certain variables affected count. *Season* - affects count. We can see that this variable is useful in determing when your seasonal peaks and lows will be with your Bikes. It seems that Fall and Summer are your busiest times. *Year* - does not affect count (as much as season). This only shows which you the difference quaritle ranges between 2011 and 2012. 2012 had more users on average. *Month* - This is affects count. This is a good variable to have. Like Season, this is a more granular approach to see when you have peaks and lows with users. It seems to line up with Season that May through October you have your peaks. *Holiday* - This is not a very good indicator and doesn’t affect count. It shows very little within the Data. *Working Day* - This is not a very good indicator and doesn’t affect count. It shows very little within the Data as well like Holiday. *Weather* - This seems to be a good indicator for count. This shows if the weather will be affecting users getting bikes. It seems that you have more users on average taking bikes when it is less chance of rain. (Logically that makes sense).

#### Task 4 - Q&A

What variables are included in your forward model? **count, hr, atemp, yr, weathersit, season, mnth, hum, weekday, holiday, windspeed and temp.** Comment on the quality of the model. Does the model match our intuition/common sense? **Yes, it seems that all of our non-categorical based data is showing positive fit into the model (P-Vlaue < 0.05)** Is there evidence of multicollinearity? **No, It seems that there is no evidence of mulitcollinearity. Example is windspeed is showing negative in the model. This being a non-categorical variable, you would want it to show positive but logically, the count is going to go down the windier it is outside.**

#### Task 5 - Q&A

Does this model differ from the forward model? If so, how? **No, both models seem to show the exact same output.**

#### Task 6 - Q&A

Workingday is being represented in the model as a categorical variable. It is being kicked out for fitting the model too well. I have created a faceted scatter plot to show the relationship between workingday and a few other variables. It shows that when it is a workingday the bike count is a lot less compared to a non-working day.

#### Task 7 - Q&A

I like the model represented. It is logical and it makes sense statistically. I would recommend this model. My only caution is that the user running the model shouldn’t solely rely on forward and backwards stepwise regression. A good model would use many different test.

options(tidyverse.quiet = TRUE)  
  
library(tidyverse)  
library(GGally)  
library(MASS)  
library(car)  
library(gridExtra)

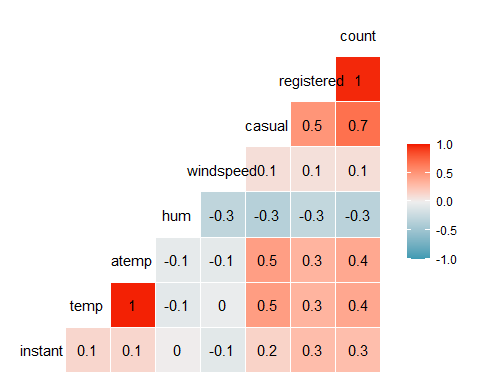
hour <- read\_csv("hour.csv")  
bike = as\_tibble(hour)  
  
bike = bike %>% mutate(season = as\_factor(as.character(season))) %>%  
mutate(season = fct\_recode(season,  
"Spring" = "1",  
"Summer" = "2",  
"Fall" = "3",  
"Winter" = "4"))  
  
bike = bike %>% mutate(yr = as\_factor(as.character(yr))) %>%  
mutate(yr = fct\_recode(yr,  
"0" = "0",  
"1" = "1"))  
  
  
# We decided to convert yr, mnth and hr variables into factors to convert them to categorical variables.   
# These values are not logically quantitative. Even though they are numbers they represent a categorical/string value.   
# We should not be running models on "Tuesday".   
  
bike = bike %>% mutate(mnth = as\_factor(as.character(mnth))) %>%  
mutate(mnth = fct\_recode(mnth,  
"1" = "1",  
"2" = "2",  
"3" = "3",  
"4" = "4",  
"5" = "5",  
"6" = "6",  
"7" = "7",  
"8" = "8",  
"9" = "9",  
"10" = "10",  
"11" = "11",  
"12" = "12"))  
  
bike = bike %>% mutate(hr = as\_factor(as.character(hr))) %>%  
mutate(hr = fct\_recode(hr,  
"0" = "0",  
"1" = "1",  
"2" = "2",  
"3" = "3",  
"4" = "4",  
"5" = "5",  
"6" = "6",  
"7" = "7",  
"8" = "8",  
"9" = "9",  
"10" = "10",  
"11" = "11",  
"12" = "12",  
"13" = "13",  
"14" = "14",  
"15" = "15",  
"16" = "16",  
"17" = "17",  
"18" = "18",  
"19" = "19",  
"20" = "20",  
"21" = "21",  
"22" = "22",  
"23" = "23"))  
  
bike = bike %>% mutate(holiday = as\_factor(as.character(holiday))) %>%  
mutate(holiday = fct\_recode(holiday,  
"NotHoliday" = "0",  
"Holiday" = "1"))  
  
bike = bike %>% mutate(workingday = as\_factor(as.character(workingday))) %>%  
mutate(workingday = fct\_recode(workingday,  
"NotWorkingDay" = "0",  
"WorkingDay" = "1"))  
  
bike = bike %>% mutate(weathersit = as\_factor(as.character(weathersit))) %>%  
mutate(weathersit = fct\_recode(weathersit,  
"NoPrecip" = "1",  
"Misty" = "2",  
"LightPrecip" = "3",  
"HeavyPrecip" = "4"))  
  
bike = bike %>% mutate(weekday = as\_factor(as.character(weekday))) %>%  
mutate(weekday = fct\_recode(weekday,  
"Sunday" = "0",  
"Monday" = "1",  
"Tuesday" = "2",  
"Wednesday" = "3",   
"Thursday" = "4",  
"Friday" = "5",  
"Saturday" = "6"))

model1 = lm(count ~ temp + atemp + hum + windspeed, bike)#linear model with quantitative variables as dependent variables, Count being independent vatiable.   
  
summary(model1)

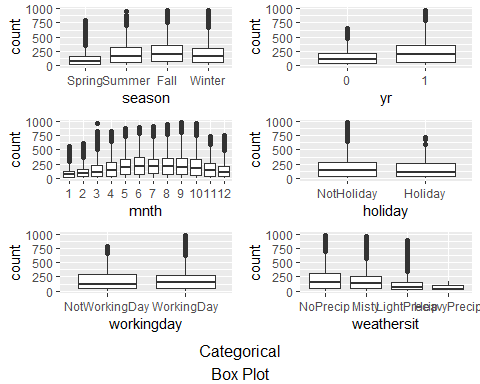
##   
## Call:  
## lm(formula = count ~ temp + atemp + hum + windspeed, data = bike)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -338.41 -101.49 -33.08 65.18 704.70   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 161.807 6.506 24.872 < 2e-16 \*\*\*  
## temp 85.576 40.751 2.100 0.0357 \*   
## atemp 314.343 45.714 6.876 6.35e-12 \*\*\*  
## hum -275.180 6.466 -42.560 < 2e-16 \*\*\*  
## windspeed 42.979 10.451 4.112 3.93e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 156.7 on 17374 degrees of freedom  
## Multiple R-squared: 0.2534, Adjusted R-squared: 0.2532   
## F-statistic: 1474 on 4 and 17374 DF, p-value: < 2.2e-16

ggcorr(bike, label = TRUE)

## Warning in ggcorr(bike, label = TRUE): data in column(s) 'dteday', 'season',  
## 'yr', 'mnth', 'hr', 'holiday', 'weekday', 'workingday', 'weathersit' are not  
## numeric and were ignored



ssn = ggplot(bike, aes(x=season, y=count)) +   
 geom\_boxplot()  
  
year = ggplot(bike, aes(x=yr, y=count)) +   
 geom\_boxplot()  
  
month = ggplot(bike, aes(x=mnth, y=count)) +   
 geom\_boxplot()  
  
hldy = ggplot(bike, aes(x=holiday, y=count)) +   
 geom\_boxplot()  
  
wkngday = ggplot(bike, aes(x=workingday, y=count)) +   
 geom\_boxplot()  
  
wethsit = ggplot(bike, aes(x=weathersit, y=count)) +   
 geom\_boxplot()  
  
  
grid.arrange(ssn, year, month, hldy, wkngday, wethsit, bottom="Categorical\nBox Plot")



bike = bike %>% dplyr::select(-c("instant", "dteday", "registered", "casual"))  
  
allmod = lm(count ~., bike)  
summary(allmod)

##   
## Call:  
## lm(formula = count ~ ., data = bike)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -393.87 -60.66 -7.96 51.31 439.18   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -67.542 6.612 -10.216 < 2e-16 \*\*\*  
## seasonSummer 38.178 4.856 7.862 4.00e-15 \*\*\*  
## seasonFall 32.055 5.749 5.575 2.51e-08 \*\*\*  
## seasonWinter 67.994 4.882 13.928 < 2e-16 \*\*\*  
## yr1 85.431 1.563 54.658 < 2e-16 \*\*\*  
## mnth2 3.426 3.920 0.874 0.38219   
## mnth3 14.299 4.407 3.244 0.00118 \*\*   
## mnth4 6.230 6.548 0.951 0.34144   
## mnth5 20.657 7.007 2.948 0.00320 \*\*   
## mnth6 6.238 7.205 0.866 0.38662   
## mnth7 -13.269 8.082 -1.642 0.10065   
## mnth8 7.897 7.879 1.002 0.31622   
## mnth9 32.269 7.001 4.609 4.07e-06 \*\*\*  
## mnth10 15.843 6.483 2.444 0.01455 \*   
## mnth11 -9.840 6.238 -1.577 0.11474   
## mnth12 -6.256 4.954 -1.263 0.20672   
## hr1 -17.294 5.345 -3.236 0.00122 \*\*   
## hr2 -26.369 5.364 -4.916 8.91e-07 \*\*\*  
## hr3 -37.112 5.403 -6.869 6.67e-12 \*\*\*  
## hr4 -40.263 5.408 -7.445 1.01e-13 \*\*\*  
## hr5 -23.501 5.373 -4.374 1.23e-05 \*\*\*  
## hr6 35.393 5.359 6.605 4.10e-11 \*\*\*  
## hr7 170.418 5.348 31.864 < 2e-16 \*\*\*  
## hr8 310.801 5.342 58.183 < 2e-16 \*\*\*  
## hr9 163.101 5.347 30.501 < 2e-16 \*\*\*  
## hr10 108.444 5.370 20.196 < 2e-16 \*\*\*  
## hr11 133.843 5.409 24.742 < 2e-16 \*\*\*  
## hr12 173.142 5.456 31.735 < 2e-16 \*\*\*  
## hr13 168.102 5.494 30.600 < 2e-16 \*\*\*  
## hr14 152.249 5.525 27.558 < 2e-16 \*\*\*  
## hr15 161.707 5.535 29.213 < 2e-16 \*\*\*  
## hr16 223.834 5.524 40.522 < 2e-16 \*\*\*  
## hr17 377.535 5.491 68.750 < 2e-16 \*\*\*  
## hr18 345.587 5.455 63.350 < 2e-16 \*\*\*  
## hr19 236.919 5.404 43.841 < 2e-16 \*\*\*  
## hr20 157.293 5.375 29.266 < 2e-16 \*\*\*  
## hr21 107.840 5.353 20.147 < 2e-16 \*\*\*  
## hr22 70.907 5.343 13.272 < 2e-16 \*\*\*  
## hr23 32.112 5.338 6.015 1.83e-09 \*\*\*  
## holidayHoliday -26.228 4.881 -5.374 7.81e-08 \*\*\*  
## weekdaySunday -16.089 2.878 -5.591 2.30e-08 \*\*\*  
## weekdayMonday -6.814 2.970 -2.294 0.02180 \*   
## weekdayTuesday -5.240 2.899 -1.807 0.07071 .   
## weekdayWednesday -2.464 2.894 -0.851 0.39469   
## weekdayThursday -2.940 2.892 -1.016 0.30947   
## weekdayFriday 1.356 2.885 0.470 0.63823   
## workingdayWorkingDay NA NA NA NA   
## weathersitMisty -10.409 1.920 -5.421 6.00e-08 \*\*\*  
## weathersitLightPrecip -65.189 3.236 -20.145 < 2e-16 \*\*\*  
## weathersitHeavyPrecip -62.580 58.893 -1.063 0.28797   
## temp 116.384 29.513 3.943 8.06e-05 \*\*\*  
## atemp 127.975 30.624 4.179 2.94e-05 \*\*\*  
## hum -82.802 5.554 -14.909 < 2e-16 \*\*\*  
## windspeed -29.167 7.052 -4.136 3.55e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 101.7 on 17326 degrees of freedom  
## Multiple R-squared: 0.6864, Adjusted R-squared: 0.6854   
## F-statistic: 729.1 on 52 and 17326 DF, p-value: < 2.2e-16

emptymod = lm(count ~ -1, bike)  
summary(emptymod)

##   
## Call:  
## lm(formula = count ~ -1, data = bike)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## 1 40 142 281 977   
##   
## No Coefficients  
##   
## Residual standard error: 262.3 on 17379 degrees of freedom

### Forward Stepwise

forwardmod = stepAIC(emptymod, direction = "forward", scope=list(upper=allmod, lower=emptymod),  
 trace = TRUE)

## Start: AIC=193582.9  
## count ~ -1  
##   
## Df Sum of Sq RSS AIC  
## + hr 23 908466780 287135941 168839  
## + temp 1 717518886 478083835 177655  
## + atemp 1 715465707 480137014 177729  
## + season 4 661570488 534032233 179584  
## + mnth 11 654010088 541592633 179843  
## + weekday 6 533645489 661957232 183320  
## + windspeed 1 493276312 702326409 184339  
## + yr 1 480966542 714636179 184641  
## + hum 1 466692801 728909920 184985  
## + workingday 1 442911387 752691334 185543  
## + weathersit 3 157106360 1038496361 191141  
## + holiday 1 12304098 1183298623 193405  
## <none> 1195602721 193583  
##   
## Step: AIC=168712.5  
## count ~ hr - 1  
##   
## Df Sum of Sq RSS AIC  
## + atemp 1 50518941 234507969 165324  
## + temp 1 50101685 234925225 165355  
## + mnth 11 44822160 240204750 165761  
## + season 3 39619754 245407156 166117  
## + yr 1 36875130 248151780 166307  
## + weathersit 3 13766672 271260238 167858  
## + hum 1 4924310 280102600 168412  
## + windspeed 1 1476211 283550699 168624  
## + holiday 1 561784 284465126 168680  
## + weekday 6 719530 284307380 168681  
## + workingday 1 485366 284541544 168685  
## <none> 285026910 168713  
##   
## Step: AIC=165324  
## count ~ hr + atemp - 1  
##   
## Df Sum of Sq RSS AIC  
## + yr 1 33463769 201044200 162650  
## + weathersit 3 9227265 225280704 164632  
## + hum 1 7008684 227499285 164799  
## + season 3 6580442 227927527 164835  
## + mnth 11 5854560 228653409 164907  
## + weekday 6 607638 233900331 165291  
## + holiday 1 274006 234233963 165306  
## + temp 1 152153 234355816 165315  
## + windspeed 1 120557 234387412 165317  
## + workingday 1 90170 234417799 165319  
## <none> 234507969 165324  
##   
## Step: AIC=162650.2  
## count ~ hr + atemp + yr - 1  
##   
## Df Sum of Sq RSS AIC  
## + weathersit 3 8408358 192635842 161914  
## + season 3 7190305 193853896 162023  
## + mnth 11 6486062 194558138 162102  
## + hum 1 4341837 196702363 162273  
## + weekday 6 641648 200402552 162607  
## + holiday 1 324763 200719438 162624  
## + windspeed 1 109311 200934889 162643  
## + workingday 1 106404 200937797 162643  
## + temp 1 91735 200952465 162644  
## <none> 201044200 162650  
##   
## Step: AIC=161913.7  
## count ~ hr + atemp + yr + weathersit - 1  
##   
## Df Sum of Sq RSS AIC  
## + season 3 7771024 184864818 161204  
## + mnth 11 7464989 185170852 161249  
## + hum 1 805099 191830743 161843  
## + weekday 6 686172 191949670 161864  
## + holiday 1 413536 192222305 161878  
## + workingday 1 212428 192423414 161897  
## + temp 1 134482 192501360 161904  
## + windspeed 1 44407 192591435 161912  
## <none> 192635842 161914  
##   
## Step: AIC=161204.1  
## count ~ hr + atemp + yr + weathersit + season - 1  
##   
## Df Sum of Sq RSS AIC  
## + mnth 11 2051323 182813495 161032  
## + hum 1 1810161 183054657 161035  
## + weekday 6 704303 184160515 161150  
## + holiday 1 392702 184472116 161169  
## + temp 1 352584 184512234 161173  
## + workingday 1 214973 184649845 161186  
## <none> 184864818 161204  
## + windspeed 1 158 184864660 161206  
##   
## Step: AIC=161032.2  
## count ~ hr + atemp + yr + weathersit + season + mnth - 1  
##   
## Df Sum of Sq RSS AIC  
## + hum 1 2356411 180457084 160809  
## + weekday 6 692672 182120823 160978  
## + holiday 1 312321 182501174 161004  
## + temp 1 233052 182580443 161012  
## + workingday 1 203953 182609542 161015  
## <none> 182813495 161032  
## + windspeed 1 68 182813428 161034  
##   
## Step: AIC=160808.7  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum -   
## 1  
##   
## Df Sum of Sq RSS AIC  
## + weekday 6 581105 179875980 160765  
## + holiday 1 322997 180134087 160780  
## + workingday 1 194139 180262945 160792  
## + windspeed 1 114287 180342797 160800  
## + temp 1 100025 180357059 160801  
## <none> 180457084 160809  
##   
## Step: AIC=160764.7  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday - 1  
##   
## Df Sum of Sq RSS AIC  
## + holiday 1 274717 179601263 160740  
## + workingday 1 274717 179601263 160740  
## + windspeed 1 112085 179763895 160756  
## + temp 1 77171 179798809 160759  
## <none> 179875980 160765  
##   
## Step: AIC=160740.1  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday + holiday - 1  
##   
## Df Sum of Sq RSS AIC  
## + windspeed 1 111562 179489701 160731  
## + temp 1 95460 179505803 160733  
## <none> 179601263 160740  
##   
## Step: AIC=160731.3  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday + holiday + windspeed - 1  
##   
## Df Sum of Sq RSS AIC  
## + temp 1 160954 179328746 160718  
## <none> 179489701 160731  
##   
## Step: AIC=160717.7  
## count ~ hr + atemp + yr + weathersit + season + mnth + hum +   
## weekday + holiday + windspeed + temp - 1  
##   
## Df Sum of Sq RSS AIC  
## <none> 179328746 160718

summary(forwardmod)

##   
## Call:  
## lm(formula = count ~ hr + atemp + yr + weathersit + season +   
## mnth + hum + weekday + holiday + windspeed + temp - 1, data = bike)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -393.87 -60.66 -7.96 51.31 439.18   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## hr0 -67.542 6.612 -10.216 < 2e-16 \*\*\*  
## hr1 -84.836 6.627 -12.801 < 2e-16 \*\*\*  
## hr2 -93.910 6.653 -14.115 < 2e-16 \*\*\*  
## hr3 -104.654 6.710 -15.596 < 2e-16 \*\*\*  
## hr4 -107.804 6.736 -16.004 < 2e-16 \*\*\*  
## hr5 -91.043 6.691 -13.608 < 2e-16 \*\*\*  
## hr6 -32.149 6.684 -4.810 1.52e-06 \*\*\*  
## hr7 102.876 6.654 15.460 < 2e-16 \*\*\*  
## hr8 243.259 6.639 36.642 < 2e-16 \*\*\*  
## hr9 95.559 6.600 14.479 < 2e-16 \*\*\*  
## hr10 40.903 6.557 6.238 4.54e-10 \*\*\*  
## hr11 66.301 6.522 10.166 < 2e-16 \*\*\*  
## hr12 105.600 6.509 16.224 < 2e-16 \*\*\*  
## hr13 100.560 6.513 15.440 < 2e-16 \*\*\*  
## hr14 84.707 6.543 12.947 < 2e-16 \*\*\*  
## hr15 94.166 6.548 14.381 < 2e-16 \*\*\*  
## hr16 156.292 6.546 23.875 < 2e-16 \*\*\*  
## hr17 309.994 6.534 47.444 < 2e-16 \*\*\*  
## hr18 278.045 6.532 42.568 < 2e-16 \*\*\*  
## hr19 169.377 6.542 25.890 < 2e-16 \*\*\*  
## hr20 89.751 6.540 13.724 < 2e-16 \*\*\*  
## hr21 40.298 6.558 6.145 8.18e-10 \*\*\*  
## hr22 3.365 6.574 0.512 0.60874   
## hr23 -35.429 6.576 -5.388 7.22e-08 \*\*\*  
## atemp 127.975 30.624 4.179 2.94e-05 \*\*\*  
## yr1 85.431 1.563 54.658 < 2e-16 \*\*\*  
## weathersitMisty -10.409 1.920 -5.421 6.00e-08 \*\*\*  
## weathersitLightPrecip -65.189 3.236 -20.145 < 2e-16 \*\*\*  
## weathersitHeavyPrecip -62.580 58.893 -1.063 0.28797   
## seasonSummer 38.178 4.856 7.862 4.00e-15 \*\*\*  
## seasonFall 32.055 5.749 5.575 2.51e-08 \*\*\*  
## seasonWinter 67.994 4.882 13.928 < 2e-16 \*\*\*  
## mnth2 3.426 3.920 0.874 0.38219   
## mnth3 14.299 4.407 3.244 0.00118 \*\*   
## mnth4 6.230 6.548 0.951 0.34144   
## mnth5 20.657 7.007 2.948 0.00320 \*\*   
## mnth6 6.238 7.205 0.866 0.38662   
## mnth7 -13.269 8.082 -1.642 0.10065   
## mnth8 7.897 7.879 1.002 0.31622   
## mnth9 32.269 7.001 4.609 4.07e-06 \*\*\*  
## mnth10 15.843 6.483 2.444 0.01455 \*   
## mnth11 -9.840 6.238 -1.577 0.11474   
## mnth12 -6.256 4.954 -1.263 0.20672   
## hum -82.802 5.554 -14.909 < 2e-16 \*\*\*  
## weekdaySunday -16.089 2.878 -5.591 2.30e-08 \*\*\*  
## weekdayMonday -6.814 2.970 -2.294 0.02180 \*   
## weekdayTuesday -5.240 2.899 -1.807 0.07071 .   
## weekdayWednesday -2.464 2.894 -0.851 0.39469   
## weekdayThursday -2.940 2.892 -1.016 0.30947   
## weekdayFriday 1.356 2.885 0.470 0.63823   
## holidayHoliday -26.228 4.881 -5.374 7.81e-08 \*\*\*  
## windspeed -29.167 7.052 -4.136 3.55e-05 \*\*\*  
## temp 116.384 29.513 3.943 8.06e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 101.7 on 17326 degrees of freedom  
## Multiple R-squared: 0.85, Adjusted R-squared: 0.8496   
## F-statistic: 1853 on 53 and 17326 DF, p-value: < 2.2e-16

### Backward Stepwise

backmod = stepAIC(allmod, direction = "backward", trace = TRUE)

## Start: AIC=160717.7  
## count ~ season + yr + mnth + hr + holiday + weekday + workingday +   
## weathersit + temp + atemp + hum + windspeed  
##   
##   
## Step: AIC=160717.7  
## count ~ season + yr + mnth + hr + holiday + weekday + weathersit +   
## temp + atemp + hum + windspeed  
##   
## Df Sum of Sq RSS AIC  
## <none> 179328746 160718  
## - temp 1 160954 179489701 160731  
## - windspeed 1 177057 179505803 160733  
## - atemp 1 180751 179509498 160733  
## - holiday 1 298893 179627639 160745  
## - weekday 6 498795 179827541 160754  
## - mnth 11 2426171 181754917 160929  
## - hum 1 2300667 181629413 160937  
## - season 3 2398467 181727213 160943  
## - weathersit 3 4208731 183537478 161115  
## - yr 1 30920851 210249597 163480  
## - hr 23 196741474 376070220 173542

summary(backmod)

##   
## Call:  
## lm(formula = count ~ season + yr + mnth + hr + holiday + weekday +   
## weathersit + temp + atemp + hum + windspeed, data = bike)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -393.87 -60.66 -7.96 51.31 439.18   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -67.542 6.612 -10.216 < 2e-16 \*\*\*  
## seasonSummer 38.178 4.856 7.862 4.00e-15 \*\*\*  
## seasonFall 32.055 5.749 5.575 2.51e-08 \*\*\*  
## seasonWinter 67.994 4.882 13.928 < 2e-16 \*\*\*  
## yr1 85.431 1.563 54.658 < 2e-16 \*\*\*  
## mnth2 3.426 3.920 0.874 0.38219   
## mnth3 14.299 4.407 3.244 0.00118 \*\*   
## mnth4 6.230 6.548 0.951 0.34144   
## mnth5 20.657 7.007 2.948 0.00320 \*\*   
## mnth6 6.238 7.205 0.866 0.38662   
## mnth7 -13.269 8.082 -1.642 0.10065   
## mnth8 7.897 7.879 1.002 0.31622   
## mnth9 32.269 7.001 4.609 4.07e-06 \*\*\*  
## mnth10 15.843 6.483 2.444 0.01455 \*   
## mnth11 -9.840 6.238 -1.577 0.11474   
## mnth12 -6.256 4.954 -1.263 0.20672   
## hr1 -17.294 5.345 -3.236 0.00122 \*\*   
## hr2 -26.369 5.364 -4.916 8.91e-07 \*\*\*  
## hr3 -37.112 5.403 -6.869 6.67e-12 \*\*\*  
## hr4 -40.263 5.408 -7.445 1.01e-13 \*\*\*  
## hr5 -23.501 5.373 -4.374 1.23e-05 \*\*\*  
## hr6 35.393 5.359 6.605 4.10e-11 \*\*\*  
## hr7 170.418 5.348 31.864 < 2e-16 \*\*\*  
## hr8 310.801 5.342 58.183 < 2e-16 \*\*\*  
## hr9 163.101 5.347 30.501 < 2e-16 \*\*\*  
## hr10 108.444 5.370 20.196 < 2e-16 \*\*\*  
## hr11 133.843 5.409 24.742 < 2e-16 \*\*\*  
## hr12 173.142 5.456 31.735 < 2e-16 \*\*\*  
## hr13 168.102 5.494 30.600 < 2e-16 \*\*\*  
## hr14 152.249 5.525 27.558 < 2e-16 \*\*\*  
## hr15 161.707 5.535 29.213 < 2e-16 \*\*\*  
## hr16 223.834 5.524 40.522 < 2e-16 \*\*\*  
## hr17 377.535 5.491 68.750 < 2e-16 \*\*\*  
## hr18 345.587 5.455 63.350 < 2e-16 \*\*\*  
## hr19 236.919 5.404 43.841 < 2e-16 \*\*\*  
## hr20 157.293 5.375 29.266 < 2e-16 \*\*\*  
## hr21 107.840 5.353 20.147 < 2e-16 \*\*\*  
## hr22 70.907 5.343 13.272 < 2e-16 \*\*\*  
## hr23 32.112 5.338 6.015 1.83e-09 \*\*\*  
## holidayHoliday -26.228 4.881 -5.374 7.81e-08 \*\*\*  
## weekdaySunday -16.089 2.878 -5.591 2.30e-08 \*\*\*  
## weekdayMonday -6.814 2.970 -2.294 0.02180 \*   
## weekdayTuesday -5.240 2.899 -1.807 0.07071 .   
## weekdayWednesday -2.464 2.894 -0.851 0.39469   
## weekdayThursday -2.940 2.892 -1.016 0.30947   
## weekdayFriday 1.356 2.885 0.470 0.63823   
## weathersitMisty -10.409 1.920 -5.421 6.00e-08 \*\*\*  
## weathersitLightPrecip -65.189 3.236 -20.145 < 2e-16 \*\*\*  
## weathersitHeavyPrecip -62.580 58.893 -1.063 0.28797   
## temp 116.384 29.513 3.943 8.06e-05 \*\*\*  
## atemp 127.975 30.624 4.179 2.94e-05 \*\*\*  
## hum -82.802 5.554 -14.909 < 2e-16 \*\*\*  
## windspeed -29.167 7.052 -4.136 3.55e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 101.7 on 17326 degrees of freedom  
## Multiple R-squared: 0.6864, Adjusted R-squared: 0.6854   
## F-statistic: 729.1 on 52 and 17326 DF, p-value: < 2.2e-16

ssn = ggplot(bike, aes(x=workingday, y=season)) + geom\_point() + theme\_bw()  
  
year = ggplot(bike, aes(x=workingday, y=yr)) + geom\_point() + theme\_bw()  
  
month = ggplot(bike, aes(x=workingday, y=mnth)) + geom\_point() + theme\_bw()  
  
hldy = ggplot(bike, aes(x=workingday, y=holiday)) + geom\_point() + theme\_bw()  
  
wethsit = ggplot(bike, aes(x=workingday, y=weathersit)) + geom\_point() + theme\_bw()  
  
cnt = ggplot(bike, aes(x=workingday, y=count)) + geom\_point() + theme\_bw()  
  
grid.arrange(ssn, year, month, hldy, wethsit, cnt, bottom="Correlation of\n'Workingday' Variable")

